

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Max-Planck-Gesellschaft zur Foerderung der
Wissenschaften e.V. Berlin
(B) STREET: Hofgartenstr. 2
(C) CITY: Muenchen
(E) COUNTRY: Germany
(F) POSTAL CODE (ZIP): 80539

(ii) TITLE OF INVENTION: Helicobacter pylori live vaccine

(iii) NUMBER OF SEQUENCES: 6

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1557 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

(vii) IMMEDIATE SOURCE:

- (B) CLONE: alpB

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 1..1554

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATG ACA CAA TCT CAA AAA GTA AGA TTC TTA GCC CCT TTA AGC CTA GCG
48
Met Thr Gln Ser Gln Lys Val Arg Phe Leu Ala Pro Leu Ser Leu Ala
1 5 10 15
TTA AGC TTG AGC TTC AAT CCA GTG GGC GCT GAA GAA GAT GGG GGC TTT

96
 Leu Ser Leu Ser Phe Asn Pro Val Gly Ala Glu Glu Asp Gly Gly Phe
 20 25 30
 ATG ACC TTT GGG TAT GAA TTA GGT CAG GTG GTC CAA CAA GTG AAA AAC
 144
 Met Thr Phe Gly Tyr Glu Leu Gly Gln Val Val Gln Gln Val Lys Asn
 35 40 45
 CCG GGT AAA ATC AAA GCC GAA GAA TTA GCC GGC TTG TTA AAC TCT ACC
 192
 Pro Gly Lys Ile Lys Ala Glu Glu Leu Ala Gly Leu Leu Asn Ser Thr
 50 55 60
 ACA ACA AAC AAC ACC AAT ATC AAT ATT GCA GGC ACA GGA GGC AAT GTC
 240
 Thr Thr Asn Asn Thr Asn Ile Asn Ile Ala Gly Thr Gly Gly Asn Val
 65 70 75 80
 GCC GGG ACT TTG GGC AAC CTT TTT ATG AAC CAA TTA GGC AAT TTG ATT
 288
 Ala Gly Thr Leu Gly Asn Leu Phe Met Asn Gln Leu Gly Asn Leu Ile
 85 90 95
 GAT TTG TAT CCC ACT TTG AAC ACT AGT AAT ATC ACA CAA TGT GGC ACT
 336
 Asp Leu Tyr Pro Thr Leu Asn Thr Ser Asn Ile Thr Gln Cys Gly Thr
 100 105 110
 ACT AAT AGT GGT AGT AGT AGT AGT GGT GGT GGT GCG GCC ACA GCC GCT
 384
 Thr Asn Ser Gly Ser Ser Ser Ser Gly Gly Gly Ala Ala Thr Ala Ala
 115 120 125
 GCT ACT ACT AGC AAT AAG CCT TGT TTC CAA GGT AAC CTG GAT CTT TAT
 432
 Ala Thr Thr Ser Asn Lys Pro Cys Phe Gln Gly Asn Leu Asp Leu Tyr
 130 135 140
 AGA AAA ATG GTT GAC TCT ATC AAA ACT TTG AGT CAA AAC ATC AGC AAG
 480
 Arg Lys Met Val Asp Ser Ile Lys Thr Leu Ser Gln Asn Ile Ser Lys
 145 150 155 160
 AAT ATC TTT CAA GGC AAC AAC AAC ACC ACG AGC CAA AAT CTC TCC AAC
 528
 Asn Ile Phe Gln Gly Asn Asn Asn Thr Thr Ser Gln Asn Leu Ser Asn
 165 170 175
 CAG CTC AGT GAG CTT AAC ACC GCT AGC GTT TAT TTG ACT TAC ATG AAC
 576
 Gln Leu Ser Glu Leu Asn Thr Ala Ser Val Tyr Leu Thr Tyr Met Asn
 180 185 190

96
 144
 192
 240
 288
 336
 384
 432
 480
 528
 576

Sub B1

Sub B1

TCG TTC TTA AAC GCC AAT AAC CAA GCG GGT GGG ATT TTT CAA AAC AAC
624
Ser Phe Leu Asn Ala Asn Asn Gln Ala Gly Gly Ile Phe Gln Asn Asn
195 200 205

ACT AAT CAA GCT TAT GGA AAT GGG GTT ACC GCT CAA CAA ATC GCT TAT
672
Thr Asn Gln Ala Tyr Gly Asn Gly Val Thr Ala Gln Gln Ile Ala Tyr
210 215 220

ATC CTA AAG CAA GCT TCA ATC ACT ATG GGG CCA AGC GGT GAT AGC GGT
720
Ile Leu Lys Gln Ala Ser Ile Thr Met Gly Pro Ser Gly Asp Ser Gly
225 230 235 240

GCT GCC GCA GCG TTT TTG GAT GCC GCT TTA GCG CAA CAT GTT TTC AAC
768
Ala Ala Ala Ala Phe Leu Asp Ala Ala Leu Ala Gln His Val Phe Asn
245 250 255

TCC GCT AAC GCC GGG AAC GAT TTG AGC GCT AAG GAA TTC ACT AGC TTG
816
Ser Ala Asn Ala Gly Asn Asp Leu Ser Ala Lys Glu Phe Thr Ser Leu
260 265 270

GTG CAA AAT ATC GTC AAT AAT TCT CAA AAC GCT TTA ACG CTA GCC AAC
864
Val Gln Asn Ile Val Asn Asn Ser Gln Asn Ala Leu Thr Leu Ala Asn
275 280 285

AAC GCT AAC ATC AGC AAT TCA ACA GGC TAT CAA GTG AGC TAT GGC GGG
912
Asn Ala Asn Ile Ser Asn Ser Thr Gly Tyr Gln Val Ser Tyr Gly Gly
290 295 300

AAT ATT GAT CAA GCG CGA TCT ACC CAA CTA TTA AAC AAC ACC ACA AAC
960
Asn Ile Asp Gln Ala Arg Ser Thr Gln Leu Leu Asn Asn Thr Thr Asn
305 310 315 320

ACT TTG GCT AAA GTT AGC GCT TTG AAT AAC GAG CTT AAA GCT AAC CCA
1008
Thr Leu Ala Lys Val Ser Ala Leu Asn Asn Glu Leu Lys Ala Asn Pro
325 330 335

TGG CTT GGG AAT TTT GCC GCC GGT AAC AGC TCT CAA GTG AAT GCG TTT
1056
Trp Leu Gly Asn Phe Ala Ala Gly Asn Ser Ser Gln Val Asn Ala Phe
340 345 350

AAC GGG TTT ATC ACT AAA ATC GGT TAC AAG CAA TTC TTT GGG GAA AAC
1104
Asn Gly Phe Ile Thr Lys Ile Gly Tyr Lys Gln Phe Phe Gly Glu Asn
355 360 365

~~AAG AAT GTG GGC TTA CGC TAC TAC GGC TTC TTC AGC TAT AAC GGC GCG
 1152
 Lys Asn Val Gly Leu Arg Tyr Tyr Gly Phe Phe Ser Tyr Asn Gly Ala
 370 375 380
 GGC GTG GGT AAT GGC CCT ACT TAC AAT CAA GTC AAT TTG CTC ACT TAT
 1200
 Gly Val Gly Asn Gly Pro Thr Tyr Asn Gln Val Asn Leu Leu Thr Tyr
 385 390 395 400
 GGG GTG GGG ACT GAT GTG CTT TAC AAT GTG TTT AGC CGC TCT TTT GGT
 1248
 Gly Val Gly Thr Asp Val Leu Tyr Asn Val Phe Ser Arg Ser Phe Gly
 405 410 415
 AGT AGG AGT CTT AAT GCG GGC TTC TTT GGG GGG ATC CAA CTC GCA GGG
 1296
 Ser Arg Ser Leu Asn Ala Gly Phe Phe Gly Gly Ile Gln Leu Ala Gly
 420 425 430
 GAT ACT TAC ATC AGC ACG CTA AGA AAC AGC TCT CAG CTT GCG AGC AGA
 1344
 Asp Thr Tyr Ile Ser Thr Leu Arg Asn Ser Ser Gln Leu Ala Ser Arg
 435 440 445
 CCT ACA GCG ACG AAA TTC CAA TTC TTG TTT GAT GTG GGC TTA CGC ATG
 1392
 Pro Thr Ala Thr Lys Phe Gln Phe Leu Phe Asp Val Gly Leu Arg Met
 450 455 460
 AAC TTT GGT ATC TTG AAA AAA GAC TTG AAA AGC CAT AAC CAG CAT TCT
 1440
 Asn Phe Gly Ile Leu Lys Lys Asp Leu Lys Ser His Asn Gln His Ser
 465 470 475 480
 ATA GAA ATC GGT GTG CAA ATC CCT ACG ATT TAC AAC ACT TAC TAT AAA
 1488
 Ile Glu Ile Gly Val Gln Ile Pro Thr Ile Tyr Asn Thr Tyr Tyr Lys
 485 490 495
 GCT GGC GGT GCT GAA GTG AAA TAC TTC CGC CCT TAT AGC GTG TAT TGG
 1536
 Ala Gly Gly Ala Glu Val Lys Tyr Phe Arg Pro Tyr Ser Val Tyr Trp
 500 505 510
 GTC TAT GGC TAC GCC TTC TAA
 1557
 Val Tyr Gly Tyr Ala Phe
 515~~

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 518 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Thr Gln Ser Gln Lys Val Arg Phe Leu Ala Pro Leu Ser Leu Ala
 1 5 10 15
 Leu Ser Leu Ser Phe Asn Pro Val Gly Ala Glu Glu Asp Gly Gly Phe
 20 25 30
 Met Thr Phe Gly Tyr Glu Leu Gly Gln Val Val Gln Gln Val Lys Asn
 35 40 45
 Pro Gly Lys Ile Lys Ala Glu Glu Leu Ala Gly Leu Leu Asn Ser Thr
 50 55 60
 Thr Thr Asn Asn Thr Asn Ile Asn Ile Ala Gly Thr Gly Gly Asn Val
 65 70 75 80
 Ala Gly Thr Leu Gly Asn Leu Phe Met Asn Gln Leu Gly Asn Leu Ile
 85 90 95
 Asp Leu Tyr Pro Thr Leu Asn Thr Ser Asn Ile Thr Gln Cys Gly Thr
 100 105 110
 Thr Asn Ser Gly Ser Ser Ser Ser Gly Gly Gly Ala Ala Thr Ala Ala
 115 120 125
 Ala Thr Thr Ser Asn Lys Pro Cys Phe Gln Gly Asn Leu Asp Leu Tyr
 130 135 140
 Arg Lys Met Val Asp Ser Ile Lys Thr Leu Ser Gln Asn Ile Ser Lys
 145 150 155 160
 Asn Ile Phe Gln Gly Asn Asn Asn Thr Thr Ser Gln Asn Leu Ser Asn
 165 170 175
 Gln Leu Ser Glu Leu Asn Thr Ala Ser Val Tyr Leu Thr Tyr Met Asn
 180 185 190
 Ser Phe Leu Asn Ala Asn Asn Gln Ala Gly Gly Ile Phe Gln Asn Asn
 195 200 205
 Thr Asn Gln Ala Tyr Gly Asn Gly Val Thr Ala Gln Gln Ile Ala Tyr
 210 215 220
 Ile Leu Lys Gln Ala Ser Ile Thr Met Gly Pro Ser Gly Asp Ser Gly
 225 230 235 240
 Ala Ala Ala Ala Phe Leu Asp Ala Ala Leu Ala Gln His Val Phe Asn
 245 250 255

Ser Ala Asn Ala Gly Asn Asp Leu Ser Ala Lys Glu Phe Thr Ser Leu
 260 265 270
 Val Gln Asn Ile Val Asn Asn Ser Gln Asn Ala Leu Thr Leu Ala Asn
 275 280 285
 Asn Ala Asn Ile Ser Asn Ser Thr Gly Tyr Gln Val Ser Tyr Gly Gly
 290 295 300
 Asn Ile Asp Gln Ala Arg Ser Thr Gln Leu Leu Asn Asn Thr Thr Asn
 305 310 315 320
 Thr Leu Ala Lys Val Ser Ala Leu Asn Asn Glu Leu Lys Ala Asn Pro
 325 330 335
 Trp Leu Gly Asn Phe Ala Ala Gly Asn Ser Ser Gln Val Asn Ala Phe
 340 345 350
 Asn Gly Phe Ile Thr Lys Ile Gly Tyr Lys Gln Phe Phe Gly Glu Asn
 355 360 365
 Lys Asn Val Gly Leu Arg Tyr Tyr Gly Phe Phe Ser Tyr Asn Gly Ala
 370 375 380
 Gly Val Gly Asn Gly Pro Thr Tyr Asn Gln Val Asn Leu Leu Thr Tyr
 385 390 395 400
 Gly Val Gly Thr Asp Val Leu Tyr Asn Val Phe Ser Arg Ser Phe Gly
 405 410 415
 Ser Arg Ser Leu Asn Ala Gly Phe Phe Gly Gly Ile Gln Leu Ala Gly
 420 425 430
 Asp Thr Tyr Ile Ser Thr Leu Arg Asn Ser Ser Gln Leu Ala Ser Arg
 435 440 445
 Pro Thr Ala Thr Lys Phe Gln Phe Leu Phe Asp Val Gly Leu Arg Met
 450 455 460
 Asn Phe Gly Ile Leu Lys Lys Asp Leu Lys Ser His Asn Gln His Ser
 465 470 475 480
 Ile Glu Ile Gly Val Gln Ile Pro Thr Ile Tyr Asn Thr Tyr Tyr Lys
 485 490 495
 Ala Gly Gly Ala Glu Val Lys Tyr Phe Arg Pro Tyr Ser Val Tyr Trp
 500 505 510
 Val Tyr Gly Tyr Ala Phe
 515

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

ATG 48	ATA	AAA	AAG	AAT	AGA	ACG	CTG	TTT	CTT	AGT	CTA	GCC	CTT	TGC	GCT
Met	Ile	Lys	Lys	Asn	Arg	Thr	Leu	Phe	Leu	Ser	Leu	Ala	Leu	Cys	Ala
520						525					530				
AGC 96	ATA	AGT	TAT	GCC	GAA	GAT	GAT	GGA	GGG	TTT	TTC	ACC	GTC	GGT	TAT
Ser	Ile	Ser	Tyr	Ala	Glu	Asp	Asp	Gly	Gly	Phe	Phe	Thr	Val	Gly	Tyr
535					540					545					550
CAG 144	CTC	GGG	CAA	GTC	ATG	CAA	GAT	GTC	CAA	AAC	CCA	GGC	GGC	GCT	AAA
Gln	Leu	Gly	Gln	Val	Met	Gln	Asp	Val	Gln	Asn	Pro	Gly	Gly	Ala	Lys
				555					560					565	
AGC 192	GAC	GAA	CTC	GCC	AGA	GAG	CTT	AAC	GCT	GAT	GTA	ACG	AAC	AAC	ATT
Ser	Asp	Glu	Leu	Ala	Arg	Glu	Leu	Asn	Ala	Asp	Val	Thr	Asn	Asn	Ile
			570					575					580		
TTA 240	AAC	AAC	AAC	ACC	GGA	GGC	AAC	ATC	GCA	GGG	GCG	TTG	AGT	AAC	GCT
Leu	Asn	Asn	Asn	Thr	Gly	Gly	Asn	Ile	Ala	Gly	Ala	Leu	Ser	Asn	Ala
		585					590					595			
TTC 288	TCC	CAA	TAC	CTT	TAT	TCG	CTT	TTA	GGG	GCT	TAC	CCC	ACA	AAA	CTC
Phe	Ser	Gln	Tyr	Leu	Tyr	Ser	Leu	Leu	Gly	Ala	Tyr	Pro	Thr	Lys	Leu
600						605					610				
AAT 336	GGT	AGC	GAT	GTG	TCT	GCG	AAC	GCT	CTT	TTA	AGT	GGT	GCG	GTA	GCG
Asn	Gly	Ser	Asp	Val	Ser	Ala	Asn	Ala	Leu	Leu	Ser	Gly	Ala	Val	Gly
615					620					625					630

TCT GGG ACT TGT GCG GCT GCA GGG ACG GCT GGT GGC ACT TCT CTT AAC
 384
 Ser Gly Thr Cys Ala Ala Ala Gly Thr Ala Gly Gly Thr Ser Leu Asn
 635 640 645
 ACT CAA AGC ACT TGC ACC GTT GCG GGC TAT TAC TGG CTC CCT AGC TTG
 432
 Thr Gln Ser Thr Cys Thr Val Ala Gly Tyr Tyr Trp Leu Pro Ser Leu
 650 655 660
 ACT GAC AGG ATT TTA AGC ACG ATC GGC AGC CAG ACT AAC TAC GGC ACG
 480
 Thr Asp Arg Ile Leu Ser Thr Ile Gly Ser Gln Thr Asn Tyr Gly Thr
 665 670 675
 AAC ACC AAT TTC CCC AAC ATG CAA CAA CAG CTC ACC TAC TTG AAT GCG
 528
 Asn Thr Asn Phe Pro Asn Met Gln Gln Gln Leu Thr Tyr Leu Asn Ala
 680 685 690
 GGG AAT GTG TTT TTT AAT GCG ATG AAT AAG GCT TTA GAG AAT AAG AAT
 576
 Gly Asn Val Phe Phe Asn Ala Met Asn Lys Ala Leu Glu Asn Lys Asn
 695 700 705 710
 GGA ACT AGT AGT GCT AGT GGA ACT AGT GGT GCG ACT GGT TCA GAT GGT
 624
 Gly Thr Ser Ser Ala Ser Gly Thr Ser Gly Ala Thr Gly Ser Asp Gly
 715 720 725
 CAA ACT TAC TCC ACA CAA GCT ATC CAA TAC CTT CAA GGC CAA CAA AAT
 672
 Gln Thr Tyr Ser Thr Gln Ala Ile Gln Tyr Leu Gln Gly Gln Gln Asn
 730 735 740
 ATC TTA AAT AAC GCA GCG AAC TTG CTC AAG CAA GAT GAA TTG CTC TTA
 720
 Ile Leu Asn Asn Ala Ala Asn Leu Leu Lys Gln Asp Glu Leu Leu Leu
 745 750 755
 GAA GCT TTC AAC TCT GCC GTA GCC GCC AAC ATT GGG AAT AAG GAA TTC
 768
 Glu Ala Phe Asn Ser Ala Val Ala Ala Asn Ile Gly Asn Lys Glu Phe
 760 765 770
 AAT TCA GCC GCT TTT ACA GGT TTG GTG CAA GGC ATT ATT GAT CAA TCT
 816
 Asn Ser Ala Ala Phe Thr Gly Leu Val Gln Gly Ile Ile Asp Gln Ser
 775 780 785 790
 CAA GCG GTT TAT AAC GAG CTC ACT AAA AAC ACC ATT AGC GGG AGT GCG
 864
 Gln Ala Val Tyr Asn Glu Leu Thr Lys Asn Thr Ile Ser Gly Ser Ala
 795 800 805

SUBB1

 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Sub 1
GTT ATT AGC GCT GGG ATA AAC TCC AAC CAA GCT AAC GCT GTG CAA GGG
912
Val Ile Ser Ala Gly Ile Asn Ser Asn Gln Ala Asn Ala Val Gln Gly
810 815 820

CGC GCT AGT CAG CTC CCT AAC GCT CTT TAT AAC GCG CAA GTA ACT TTG
960
Arg Ala Ser Gln Leu Pro Asn Ala Leu Tyr Asn Ala Gln Val Thr Leu
825 830 835

GAT AAA ATC AAT GCG CTC AAT AAT CAA GTG AGA AGC ATG CCT TAC TTG
1008
Asp Lys Ile Asn Ala Leu Asn Asn Gln Val Arg Ser Met Pro Tyr Leu
840 845 850

CCC CAA TTC AGA GCC GGG AAC AGC CGT TCA ACG AAT ATT TTA AAC GGG
1056
Pro Gln Phe Arg Ala Gly Asn Ser Arg Ser Thr Asn Ile Leu Asn Gly
855 860 865 870

TTT TAC ACC AAA ATA GGC TAT AAG CAA TTC TTC GGG AAG AAA AGG AAT
1104
Phe Tyr Thr Lys Ile Gly Tyr Lys Gln Phe Phe Gly Lys Lys Arg Asn
875 880 885

ATC GGT TTG CGC TAT TAT GGT TTC TTT TCT TAT AAC GGA GCG AGC GTG
1152
Ile Gly Leu Arg Tyr Tyr Gly Phe Phe Ser Tyr Asn Gly Ala Ser Val
890 895 900

GGC TTT AGA TCC ACT CAA AAT AAT GTA GGG TTA TAC ACT TAT GGG GTG
1200
Gly Phe Arg Ser Thr Gln Asn Asn Val Gly Leu Tyr Thr Tyr Gly Val
905 910 915

GGG ACT GAT GTG TTG TAT AAC ATC TTT AGC CGC TCC TAT CAA AAC CGC
1248
Gly Thr Asp Val Leu Tyr Asn Ile Phe Ser Arg Ser Tyr Gln Asn Arg
920 925 930

TCT GTG GAT ATG GGC TTT TTT AGC GGT ATC CAA TTA GCC GGT GAG ACC
1296
Ser Val Asp Met Gly Phe Phe Ser Gly Ile Gln Leu Ala Gly Glu Thr
935 940 945 950

TTC CAA TCC ACG CTC AGA GAT GAC CCC AAT GTG AAA TTG CAT GGG AAA
1344
Phe Gln Ser Thr Leu Arg Asp Asp Pro Asn Val Lys Leu His Gly Lys
955 960 965

ATC AAT AAC ACG CAC TTC CAG TTC CTC TTT GAC TTC GGT ATG AGG ATG
1392
Ile Asn Asn Thr His Phe Gln Phe Leu Phe Asp Phe Gly Met Arg Met
970 975 980

~~AAC TTC GGT AAG TTG GAC GGG AAA TCC AAC CGC CAC AAC CAG CAC ACG
 1440
 Asn Phe Gly Lys Leu Asp Gly Lys Ser Asn Arg His Asn Gln His Thr
 985 990 995
 GTG GAA TTT GGC GTA GTG GTG CCT ACG ATT TAT AAC ACT TAT TAC AAA
 1488
 Val Glu Phe Gly Val Val Val Pro Thr Ile Tyr Asn Thr Tyr Tyr Lys
 1000 1005 1010
 TCA GCA GGG ACT ACC GTG AAG TAT TTC CGT CCT TAT AGC GTT TAT TGG
 1536
 Ser Ala Gly Thr Thr Val Lys Tyr Phe Arg Pro Tyr Ser Val Tyr Trp
 1015 1020 1025 1030
 TCT TAT GGG TAT TCA TTC TAA
 1557
 Ser Tyr Gly Tyr Ser Phe
 1035~~

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 518 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

~~Met Ile Lys Lys Asn Arg Thr Leu Phe Leu Ser Leu Ala Leu Cys Ala
 1 5 10 15
 Ser Ile Ser Tyr Ala Glu Asp Asp Gly Gly Phe Phe Thr Val Gly Tyr
 20 25 30
 Gln Leu Gly Gln Val Met Gln Asp Val Gln Asn Pro Gly Gly Ala Lys
 35 40 45
 Ser Asp Glu Leu Ala Arg Glu Leu Asn Ala Asp Val Thr Asn Asn Ile
 50 55 60
 Leu Asn Asn Asn Thr Gly Gly Asn Ile Ala Gly Ala Leu Ser Asn Ala
 65 70 75 80
 Phe Ser Gln Tyr Leu Tyr Ser Leu Leu Gly Ala Tyr Pro Thr Lys Leu
 85 90 95
 Asn Gly Ser Asp Val Ser Ala Asn Ala Leu Leu Ser Gly Ala Val Gly
 100 105 110
 Ser Gly Thr Cys Ala Ala Ala Gly Thr Ala Gly Gly Thr Ser Leu Asn
 115 120 125~~

~~Thr Gln Ser Thr Cys Thr Val Ala Gly Tyr Tyr Trp Leu Pro Ser Leu
 130 135 140
 Thr Asp Arg Ile Leu Ser Thr Ile Gly Ser Gln Thr Asn Tyr Gly Thr
 145 150 155 160
 Asn Thr Asn Phe Pro Asn Met Gln Gln Gln Leu Thr Tyr Leu Asn Ala
 165 170 175
 Gly Asn Val Phe Phe Asn Ala Met Asn Lys Ala Leu Glu Asn Lys Asn
 180 185 190
 Gly Thr Ser Ser Ala Ser Gly Thr Ser Gly Ala Thr Gly Ser Asp Gly
 195 200 205
 Gln Thr Tyr Ser Thr Gln Ala Ile Gln Tyr Leu Gln Gly Gln Gln Asn
 210 215 220
 Ile Leu Asn Asn Ala Ala Asn Leu Leu Lys Gln Asp Glu Leu Leu Leu
 225 230 235 240
 Glu Ala Phe Asn Ser Ala Val Ala Ala Asn Ile Gly Asn Lys Glu Phe
 245 250 255
 Asn Ser Ala Ala Phe Thr Gly Leu Val Gln Gly Ile Ile Asp Gln Ser
 260 265 270
 Gln Ala Val Tyr Asn Glu Leu Thr Lys Asn Thr Ile Ser Gly Ser Ala
 275 280 285
 Val Ile Ser Ala Gly Ile Asn Ser Asn Gln Ala Asn Ala Val Gln Gly
 290 295 300
 Arg Ala Ser Gln Leu Pro Asn Ala Leu Tyr Asn Ala Gln Val Thr Leu
 305 310 315 320
 Asp Lys Ile Asn Ala Leu Asn Asn Gln Val Arg Ser Met Pro Tyr Leu
 325 330 335
 Pro Gln Phe Arg Ala Gly Asn Ser Arg Ser Thr Asn Ile Leu Asn Gly
 340 345 350
 Phe Tyr Thr Lys Ile Gly Tyr Lys Gln Phe Phe Gly Lys Lys Arg Asn
 355 360 365
 Ile Gly Leu Arg Tyr Tyr Gly Phe Phe Ser Tyr Asn Gly Ala Ser Val
 370 375 380
 Gly Phe Arg Ser Thr Gln Asn Asn Val Gly Leu Tyr Thr Tyr Gly Val
 385 390 395 400
 Gly Thr Asp Val Leu Tyr Asn Ile Phe Ser Arg Ser Tyr Gln Asn Arg
 405 410 415~~

Sub B1

Ser	Val	Asp	Met	Gly	Phe	Phe	Ser	Gly	Ile	Gln	Leu	Ala	Gly	Glu	Thr
			420					425					430		
Phe	Gln	Ser	Thr	Leu	Arg	Asp	Asp	Pro	Asn	Val	Lys	Leu	His	Gly	Lys
		435				440						445			
Ile	Asn	Asn	Thr	His	Phe	Gln	Phe	Leu	Phe	Asp	Phe	Gly	Met	Arg	Met
	450					455					460				
Asn	Phe	Gly	Lys	Leu	Asp	Gly	Lys	Ser	Asn	Arg	His	Asn	Gln	His	Thr
465					470					475					480
Val	Glu	Phe	Gly	Val	Val	Val	Pro	Thr	Ile	Tyr	Asn	Thr	Tyr	Tyr	Lys
				485					490					495	
Ser	Ala	Gly	Thr	Thr	Val	Lys	Tyr	Phe	Arg	Pro	Tyr	Ser	Val	Tyr	Trp
			500					505					510		
Ser	Tyr	Gly	Tyr	Ser	Phe										
		515													

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 656 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: both

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 567..656

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AGATCTATGA ATCTATGATA TCAACACTCT TTTTGATAAA TTTTCTCGAG GTACCGAGCT
60

TGAGGCATCA AATAAAACGA AAGGCTCAGT CGAAAGACTG GGCCTTTCGT TTTATCTGTT
120

GTTTGTCGGT GAACGCTCTC CTGAGTAGGA CAAATCCGCC GGGAGCGGAT TTGAACGTTG
180

CGAAGCAACG GCCCGGAGGG TGGCGGGCAG GACGCCCGCC ATAAACTGCC ACAAGCTCGG
240

TACCGTTGAT CTTCTATGG TGCACCTCTCA GTACAATCTG CTCTGATGCG CTACGTGACT
300

GGGTCATGGC TGCGCCCCGA CACCCGCCAA CACCCGCTGA CGCGCCCTGA CGGGCTTGTC

Sub B1
360TGCTCCCGGC ATCCGCTTAC AGACAAGCTG TGACCGTCTC CGGGAGCTGC ATGTGTCAGA
420GGTTTTTCACC GTCATCACCG AAACGCGCGA GGCCCAGCGC TTCGAACTTC TGATAGACTT
480CGAAATTAAT ACGACTCACT ATAGGGAGAC CACAACGGTT TCCCTCTAGA AATAATTTTG
540TTTAACTTTA AGAAGGAGAT ATACAT ATG AAA CTG ACT CCC AAA GAG TTA GAC
593

Met Lys Leu Thr Pro Lys Glu Leu Asp

520

525

AAG TTG ATG CTC CAC TAC GCT GGA GAA TTG GCT AAA AAA CGC AAA GAA
641Lys Leu Met Leu His Tyr Ala Gly Glu Leu Ala Lys Lys Arg Lys Glu
530 535 540AAA GGC ATT AAG CTT
656Lys Gly Ile Lys Leu
545

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Lys Leu Thr Pro Lys Glu Leu Asp Lys Leu Met Leu His Tyr Ala
1 5 10 15Gly Glu Leu Ala Lys Lys Arg Lys Glu Lys Gly Ile Lys Leu
20 25 30